February 1963

THE ASYMPTOTIC NORMALITY OF TWO TEST STATISTICS[1/] ASSOCIATED WITH THE TWO-SAMPLE PROBLEM

Saul Blumenthal

Technical Report No. 22

University of Minnesota

Minneapolis, Minnesota

Technical Report No. 85
Applied Mathematics and Statistics Laboratories
Stanford University
Stanford, California

1/ This research was supported in part by the National Science Foundation under Grant Number G-19126 at the University of Minnesota and by the Office of Naval Research under Nonr 225(52) at Stanford University.

TABLE OF CONTENTS

THE ASYMPTOTIC NORMALITY OF TWO TEST STATISTICS ASSOCIATED WITH THE TWO-SAMPLE PROBLEM

Saul Blumenthal

0. Summary.

In this paper we prove the asymptotic normality of two statistics which have been proposed to test the hypothesis that two samples come from the same parent population. One statistic is the number of runs of X's and Y's in the combined sample of X's and Y's; the other is the sum of squares of "S_i's" where S_i is the number of X's falling between the $i^{\underline{th}}$ and $(i-1)^{\underline{st}}$ largest Y's. Both statistics have been studied previously, both lead to consistent tests, and both were known to be asymptotically normal under the null distribution. Here we prove limiting normality under a fairly wide class of alternatives. We also compare the limiting power of these tests. Our method, a study of conditional moments, can also be used to prove limiting normality of "combinatorial" statistics of greater generality than the "sum of squares" statistic which we study in detail.

1. Introduction.

The purpose of this paper is to demonstrate the asymptotic normality of certain statistics which have been proposed for testing the "two sample" problem. Chief among these are the Wald-Wolfowitz run statistic and a statistic studied by Dixon [5] and by Blum and Weiss [1]. Since previous proofs of normality under the null hypothesis exist, the main contribution here is the proof of normality under a fairly wide class of alternative distributions. Using this result power can be computed for the tests in question. A comparison of limiting powers for these tests is made in Section 8.

Let $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ be two sets of independent random variables, the first set with common c.d.f. $F(x)$ and the second set with common c.d.f. $G(x)$. We assume that both $F(x)$ and $G(x)$ are absolutely continuous, and have continuous differentiable density functions $f(x)$ and $g(x)$, respectively. Further, we must assume for the purpose of using a convergence theorem in Section 3 (Theorem 3.2) that $g(F^{-1}(x))/f(F^{-1}(x))$ $(0 \leq x \leq 1)$ is bounded. Whether this condition can be relaxed is open. For the puposes of proving a result about asymptotic normality of sample spacings in Section 7, we need to assume that $g(F^{-1}(x))/f(F^{-1}(x))$ $(0 \leq x \leq 1)$ is also bounded away from zero. By the symmetry of the problem and the arbitrariness of labeling X and Y, these two conditions imply one another. Thus if one can be relaxed, both can be. We can note that if "truncated" tests are used, most distributions will meet these conditions, however distributions commonly encountered fail to meet these two conditions when x is near 0 or 1 or both.

We assume that $(m/n) = r+r_n$ where $\sqrt{n}\, r_n \rightarrow 0$ as n increases. In the sequel, we treat m/n as a constant r without loss of generality.

Let $Z_o = G^{-1}(0)$, $Z_{n+1} = G^{-1}(1)$, and $Z_1 <\ldots< Z_n$ be the values of the Y's arranged in increasing order. For each $i=1,\ldots,n+1$, let S_i be the number of X's which lie in the interval $[Z_{i-1}, Z_i]$. All the statistics to be considered can be expressed as functions of the S_i. Since the S_i are invariant under probability transformations, we shall assume hereafter that $f(x)=1$ for $0 \leq x \leq 1$, that $G(x)$ assigns unit mass to $[0, 1]$, that $G^{-1}(0)=0$, $G^{-1}(1)=1$ and that $g(x)$ is bounded above and below (away from zero). This last assumption assures the uniqueness of the inverse $G^{-1}(x)$ for all x in $[0, 1]$.

We shall denote the difference, or sample spacing, $Z_i - Z_{i-1}$ by W_i, $i=1,\ldots,n+1$.

The statistic proposed by Dixon is

$$V^2 = \frac{1}{n}\sum_{i=1}^{n+1} S_i^2 = \frac{1}{n}\sum_{i=1}^{n+1} S_i(S_i-1) + \frac{m}{n} = \frac{1}{n}\sum_{i=1}^{n+1} S_i(S_i-1) + r.$$

In Section 3, we study the distributions of "combinatorial" statistics of the form $\frac{1}{n}\sum_{i=1}^{n+1} \binom{S_i}{k}$. Clearly, V^2 has the same limiting distribution as $r + \frac{2}{n}\sum_{i=1}^{n+1} \binom{S_i}{2}$. Further, it is obvious that as test statistics, V^2 and $\frac{1}{n}\sum_{i=1}^{n+1} \binom{S_i}{2}$ will have the same properties. One could, in fact, consider the possibility of using $\frac{1}{n}\sum_{i=1}^{n+1} \binom{S_i}{k}$ as a test statistic for k other than 2. Consistency or lack thereof can be established easily using the convergence theorem of Blum and Weiss [1], and power could be computed using the results of our Section 3. We see no point in doing this here since in [1], V^2 was shown to have some desirable power properties and no similar properties have been established for other values of k. In Section 4, the power of the test based on V^2 is written down explicitly.

The run test is studied in Section 5 and its relation to the quantities $S_1,\ldots,S_{n+1}$ is indicated there. This relation is exploited to prove the limiting normality of the run statistic by obtaining the limiting normality of a certain function of $S_1,\ldots,S_{n+1}$, namely $\frac{1}{n}\sum_{i=1}^{n+1} \delta_o(S_i)$, where $\delta_o(x) = 1$ if x=0 and 0 otherwise.

The methods of proof in Sections 3 and 5 are similar and are justified by the argument given in Section 2.

It should be mentioned that tests for the one-sample test of fit which are based on statistics analogous to the above mentioned ones were proposed and studied by David [4], Kitabatake [8] and Okomoto [9], [10]. In the one-sample case, the sample intervals are $[F_o^{-1}(\frac{i-1}{n}), F_o^{-1}(\frac{i}{n})]$ $(i=1,\ldots,n)$ where $F_o(x)$ is the hypothesized distribution. The S_i are then the numbers of X's in these intervals (now fixed instead of being random). Because of the strong resemblance of the statistics, many of the computational schemes used by Kitabatake and Okomoto can be used for the two-sample case (see Sections 3 and 5).

In a recent article, Wilks [15] considers another statistic based on $S_1,\ldots,S_{n+1}$. He also indicates the utility of the one-sample methods although he does not elaborate in much detail on how they are to be used.

2. General Approach.

In both proofs of normality (Sections 3 and 5) a conditional method of moments is used to establish the asymptotic normality given the $Y_1,\ldots,Y_n$ of a function $H(S_1,\ldots,S_{n+1})$. (In Section 3, $H(S_1,\ldots,S_n)$ is $\frac{1}{n}\sum_{i=1}^{n+1} S_i^2$, and in Section 5 it is $\frac{1}{n}\sum_{i=1}^{n+1}\delta_o(S_i)$ where $\delta_o(x) = 1$ if $x=0$ and 0 otherwise.) This normality will be shown to hold for almost every sample sequence $Y_1,Y_2,\ldots$. We now justify the particular method employed. Denote $H(S_1,\ldots,S_{n+1})$ by $H_n(S)$. Denote conditional expectation given $Y_1,\ldots,Y_n$ as $E_n(\cdot|Y)$ Our goal is to show that as n increases

$$E\, e^{it\sqrt{n}(H_n(S)-E\,H_n(S))} \longrightarrow e^{-\frac{t^2c}{2}} \quad . \tag{2.1}$$

We summarize our assumptions and result as

Theorem 2.1: If $H_n(S)$ and $E_n(H_n(S)|Y)$ are as given above, if $\sqrt{n}[E_n(H_n(S)|Y) - E\, H_n(S)]$ considered as a function of $(Y_1,\ldots,Y_n)$ has a limiting non-degenerate Normal distribution, $N(0, c_1)$ and if

$$(2.2) \qquad n^{p/2}\, E_n\{[H_n(S) - E_n(H_n(S)|Y)]^p|Y\} \longrightarrow \begin{cases} (p-1)(p-3)\ldots 3.1c_2^{p/2} & \text{if } p \text{ is even} \\ 0 & \text{if } p \text{ is odd} \end{cases}$$

with probability one, (where c_2 is some constant) then (2.1) is true with $c=c_1+c_2$.

Proof: We can rewrite the expectation in (2.1) as

$$(2.3) \qquad E\, e^{it\sqrt{n}[E_n(H_n(S)|Y) - E\, H_n(S)]}\, E_n[e^{it\sqrt{n}(H_n(S) - E_n(H_n(S)|Y))}|Y].$$

The normality proof then consists of showing that the random variable $E_n[e^{it\sqrt{n}(H_n(S) - E_n(H_n(S)|Y)}|Y]$ approaches $e^{-\frac{t^2c_2}{2}}$ with probability one as n increases, where c_2 is an appropriate constant, and of showing that

$E\, e^{it\sqrt{n}[E_n(H_n(S)|Y) - E\, H_n(S)]}$ approaches $e^{-\frac{t^2c_1}{2}}$ as n increases. This latter convergence follows easily from the Lévy uniqueness theorem for characteristic functions and from our normality assumption. Thus if we can show the above convergence with probability one, because of the boundedness in absolute value of the exponentials in the expectations in (2.3), it must be that the limit as n increases of (2.3) is

$$(2.4) \qquad e^{-\frac{t^2c_2}{2}} \lim_{n\to\infty} E\, e^{it\sqrt{n}[E_n(H_n(S)|Y) - E\, H_n(S)]},$$

which is in turn (by the result noted above) $e^{-\frac{t^2}{2}(c_1+c_2)}$, and this is the desired result. (In the cases which we shall consider, $E_n(H_n(S)|Y)$ is known to have asymptotically a normal distribution from previous work (Weiss [14], Proschan

and thus we have part
and Pyke [11]) of our work done in advance.)

To show the convergence with probability one of

$E_n\{e^{it\sqrt{n}(H_n(S) - E_n(H_n(S)|Y))}|Y\}$ we shall study in Sections 3 and 5 the behavior of the moments $E_n\{[\sqrt{n}(H_n(S) - E_n(H_n(S)|Y))]^p|Y\}$ p=1,2,3,... and in particular shall show that (2.2) holds.

A series expansion (with error term) of the expression

$E_n[e^{it\sqrt{n}(H_n(S) - E_n(H_n(S)|Y))}|Y]$ shows that the result (2.2) is sufficient to imply the desired convergence with probability one. This proves the theorem and shows the direction we follow in the sequel.

3. Normality of Combinatorial Statistics.

In this section, we shall consider the limiting distributions of statistics of the form

$$(3.1)\qquad H_n^k(S) = \frac{1}{n}\sum_{i=1}^{n+1} \frac{S_i(S_i-1)\ldots(S_i-k)}{k!} = \frac{1}{n}\sum_{i=1}^{n+1} \binom{S_i}{k}.$$

$H_n^k(S)$ has the following interpretation: Consider all $\binom{m}{k}$ k-tuples $(X_{i_1},\ldots,X_{i_k})$ $1 \leq i_1 <\ldots< i_k \leq m$ of the X's, and count the number of these such that all of $X_{i_1},\ldots,X_{i_k}$ fall in the same sample interval $[Z_{j-1}, Z_j]$, j=1,...,n+1. Although we shall carry out the details only for k=2, it will be seen that the method will suffice for any k, and in fact will suffice to show the limiting joint normality of any finite set $(H_n^{k_1}(S),\ldots,H_n^{k_p}(S))$ of p of these quantities. Noting that $H_n^1(S) = \frac{1}{n}\sum_{i=1}^{n+1} S_i = \frac{m}{n} = r$, it can then be seen that the result obtained for $H_n^k(S)$ implies the limiting normality of $\frac{1}{n}\sum_{i=1}^{n+1} S_i^k$ since the latter is a linear combination of $H_n^p(S)$, $p \le k$. The same argument shows that finite collections of the form, $\left(\frac{1}{n}\sum_{i=1}^{n+1} S_i^{k_1},\ldots,\frac{1}{n}\sum_{i=1}^{n+1} S_i^{k_p}\right)$ have a limiting joint normal distribution.

For real numbers $x_1,\ldots,x_k$ such that $0 < x_i < 1$, $(i=1,\ldots,k)$, we define

$$(3.2)\qquad t_k(x_1,\ldots,x_k) = 1 \quad \text{if } x_1,\ldots,x_k \text{ fall in the same sample interval}$$
$$= 0 \quad \text{otherwise}$$

Note that implicitly $t_k(x_1,\ldots,x_k)$ is a function of $Y_1,\ldots,Y_n$ as well as $x_1,\ldots,x_k$. Since the X's are independent, we would have that $P[t_k(X_{i_1},\ldots,X_{i_k}) = 1|Y] = \sum_{i=1}^{n+1} W_i^k$ where W_i is the length of the $i^{\underline{th}}$ sample interval (based on $Y_1,\ldots,Y_n$). Note that we can write

$$(3.3)\qquad H_n^k(S) = \frac{1}{n} \sum t_k(X_{i_1},\ldots,X_{i_k})$$

The sum Σ extends over all k-tuples $(i_1,\ldots,i_k)$ $1 \le i_1 <\ldots< i_k \le m$ unless otherwise stated. In the form (3.3), $H_n^k(S)$ looks deceptively like a "U statistic", which it is not in the strictest sense. Thus we cannot use the theorems for "U statistics" but must treat this separately. Note that

$$(3.4)\qquad E_n(t_k(X_{i_1},\ldots,X_{i_k})|Y) = \sum_{i=1}^{n+1} W_i^k$$

Thus we have that

$$(3.5)\qquad E_n(H_n^k(S)|Y) = \frac{1}{n} \sum E_n(t(X_{i_1},\ldots,X_{i_k})|Y) = \frac{1}{n} \sum \left(\sum_{i=1}^{n+1} W_i^k\right)$$
$$= \frac{\binom{m}{k}}{n}\left(\sum_{i=1}^{n+1} W_i^k\right) = n^{k-1}\frac{r^k}{k!}\left(\sum_{i=1}^{n+1} W_i^k\right) + \delta_n,$$

where $\sqrt{n}\,\delta_n$ approaches 0 stochastically as n increases (see (3.21)).

The limiting standard normality of

$$(3.6)\qquad \frac{\sqrt{n}\left(n^{k-1}\sum_{i=1}^{n+1} W_i^k - k!\int_0^1 g^{1-k}(x)dx\right)}{\{[(2k)! - 2k(k!)^2]\int_0^1 g^{1-2k}(x)dx - [(k-1)\,k!\int_0^1 g^{1-k}(x)dx]^2\}^{\frac{1}{2}}}$$

has been demonstrated by Weiss [14], and again by Proschan and Pyke [11].

In view of Theorem 2.1, it remains to study the conditional moments of

$$(3.7)\qquad \sqrt{n}\left[\frac{1}{n}\sum\ (t_k(X_{i_1},\ldots,X_{i_k}) - \sum_{i=1}^{n+1} W_i^k)\right]$$

in order to verify (2.2).

Counting the various terms involved in the moments becomes very complicated, and to avoid excessive notational troubles, we shall study in detail only the case k=2.

We shall show that

Theorem 3.1 Let $t_2(X_i, X_j)$ be defined by (3.2), and g(x) the density of the Y's be bounded on [0, 1], then

$$(3.8)\qquad \lim_{n\to\infty} E_n[\{\frac{1}{\sqrt{n}}\Sigma(t_2(X_i, X_j) - \sum_{i=1}^{n+1} W_i^2)\}^p|Y] = \begin{cases} 0 & p=1,3,5\ldots \\ [(p-1)(p-3)\ldots 3.1]c^{p/2} & p=2,4\ldots \end{cases}$$

with probability one, where Σ extends over all pairs $(i < j)$. The constant c is given by

$$c = r^2\left[\int_0^1 g^{-1}(x)dx + 6r\int_0^1 g^{-2}(x)dx - 4r\left(\int_0^1 g^{-1}(x)dx\right)^2\right].$$

Proof:

Our methods of counting in the proof of (3.8) are based on those used by Daniels [3], Hoeffding [7], and Okamoto [10], chiefly the last.

Let

$$(3.9)\qquad \psi_2(X_i, X_j) = t_2(X_i, X_j) - \sum_{i=1}^{n+1} W_i^2$$

We are studying

$$(3.10)\qquad \mu_p = E_n[n^{-p/2}(\Sigma\psi_2(X_i, X_j))^p|Y]$$

which can be written as

$$(3.11)\qquad \mu_p = n^{-p/2}\,\Sigma\, E_n[\psi_2(X_{i_1}; X_{j_1}) \dots \psi_2(X_{i_p}, X_{j_p})|Y]$$

where summation is extended over all sets of pairs $(i_1, j_1),\dots,(i_p, j_p)$,

$1 \leqq i_k < j_k \leqq m$, $k=1,\dots,p$. Following Okamoto, let d denote the number of different integers among

$$(3.12)\qquad i_1, j_1; \dots; i_p\ j_p.$$

We now divide the p pairs of subscripts into e classes. Pairs of subscripts in the same class will be "linked" in the sense used by Daniels [3], and pairs not in the same class will not be "linked". The classes of "linked" pairs are equivalence classes and the members can be found as follows. The pairs (i_k, j_k) and (i_ℓ, j_ℓ) are said to be linked neighbors if one or more of the equations $i_k=i_\ell$; $i_k=j_\ell$, $j_k=i_\ell$, $j_k=k_\ell$ are satisfied. The pairs (i_k, j_k) and (i_ℓ, j_ℓ) are linked if either they are linked neighbors or there is a pair (i_α, j_α) so that either (i_k, j_k) and (i_α, j_α) are linked neighbors and (i_α, j_α) and (i_ℓ, j_ℓ) are linked or (i_ℓ, j_ℓ) and (i_α, j_α) are linked neighbors and (i_α, j_α) and (i_k, j_k) are linked. This inductive definition uniquely determines the sub-classification of the p pairs of subscripts.

We then write

$$(3.13)\qquad \mu_p = \sum_{e=1}^{p} \sum_{d=2}^{2p} A_{ed},$$

where

$$(3.14)\qquad A_{ed} = n^{-p/2}\,\Sigma^{(e,d)} E_n[\psi_2(X_{i_1}, X_{j_1}) \dots \psi_2(X_{i_p}, X_{j_p})|Y].$$

$\Sigma^{(e,d)}$ standing for summation over all sets of pairs $(i_1,j_1),\ldots,(i_p,j_p)$ such that the number of different integers is d and the number of equivalence classes is e. We shall now investigate

(3.15) $\quad E_n[\psi_2(X_{i_1}, X_{j_1}) \cdots \psi_2(X_{i_p}, X_{j_p})|Y]$

in A_{ed}. Let e equivalence classes consist of $p_1,\ldots,p_e$ pairs. Obviously,

(3.16) $\quad p = p_1 +\ldots+ p_e$.

To evaluate (3.15), we can assume without loss of generality that these classes are (we put the subscripts in parentheses after the i's and j's to simplify typing)

(3.16.1) $\quad i(1),j(1);\ldots;i(p_1),j(p_1),$

(3.16.2) $\quad i(p_1+1),j(p_1+1);\ldots;i(p_1+p_2),j(p_1+p_2) \ldots$

(3.16.e) $\quad i(p_1+p_2+\ldots+p_{e-1}+1),j(p_1+p_2+\ldots+p_{e-1}+1);\ldots;i(p),j(p).$

By independence of the X's, we have $E_n(\cdot|Y)$ in (3.15) distributed to e classes, and (3.15) becomes the product of e expectations

(3.17.1) $\quad E_n[\psi_2(X_{i(1)},X_{j(1)}) \cdots \psi_2(X_{i(p_1)},X_{j(p_1)})|Y],$

(3.17.2) $\quad E_n[\psi_2(X_{i(p_1+1)},X_{j(p_1+1)}) \cdots \psi_2(X_{i(p_1+p_2)},X_{j(p_1+p_2)})|Y] \ldots$

(3.17.e) $\quad E_n[\psi_2(X_{i(p-p_e+1)},X_{j(p-p_e+1)}) \cdots \psi\ (X_{i(p)},X_{j(p)})|Y].$

Denoting by d_g the number of different integers in the class (3.16.g) $g=1,2,\ldots,e$ $\quad (d_g \leq p_g)$ we have

(3.18) $\quad d = d_1+d_2+\ldots+d_e$.

The conditional probability $P_n(\cdot|Y)$ that α X's fall in the same interval is $\sum_{i=1}^{n+1} W_i^{\alpha}$, so that expanding the product in (3.17.g), using the definition (3.9), we have that the expectation (3.17.g) is of the form

$$(3.19)\quad \Big(\sum_{i=1}^{n+1} W_i^{d_g}\Big) + a_{01}\Big(\sum_{i=1}^{n+1} W_i^{d_g}\Big)\Big(\sum_{i=1}^{n+1} W_i^2\Big) + a_{11}\Big(\sum_{i=1}^{n+1} W_i^{d_g-1}\Big)\Big(\sum_{i=1}^{n+1} W_i^2\Big)$$

$$+ a_{02}\Big(\sum_{i=1}^{n+1} W_i^{d_g}\Big)\Big(\sum_{i=1}^{n+1} W_i^2\Big)^2 + a_{12}\Big(\sum_{i=1}^{n+1} W_i^{d_g-1}\Big)\Big(\sum_{i=1}^{n+1} W_i^2\Big)$$

$$+ a_{22}\Big(\sum_{i=1}^{n+1} W_i^{d_g-2}\Big)\Big(\sum_{i=1}^{n+1} W_i^2\Big) + \dots + a_{d_g p_g}\Big(\sum_{i=1}^{n+1} W_i^2\Big)^{p_g}.$$

Clearly, the coefficients of the above polynomial, the a_{ij}'s depend on the d_g and p_g, and not on n. We need not evaluate them specifically. At this point, we call upon a result due to Weiss [13].

Theorem 3.2 (Weiss) For each $t \geq 0$, let $R_n(t)$ be the proportion among $W_1,\dots,W_{n+1}$ which do not exceed t/n, and let

$$(3.20)\quad R(t) = 1 - \int_0^1 e^{-tg(x)} g(x)dx.$$

Then if $g(x)$ is bounded on $[0,1]$,

$$(3.21)\quad P\{\lim_{n\to\infty} \sup_{t\geq 0} |R_n(t) - R(t)| = 0\} = 1.$$

Thus, we have

$$(3.22)\quad n^{\alpha-1}\sum_{i=1}^{n+1} W_i^{\alpha} = \Big(1+\frac{1}{n}\Big)\int_0^n t^{\alpha} dR_n(t)$$

and

$$(3.23)\quad \lim_{n\to\infty} n^{\alpha-1}\sum_{i=1}^{n+1} W_i^{\alpha} = \int_0^{\infty} t^{\alpha}\int_0^1 e^{-tg(x)} g^2(x)dxdt = \Gamma(\alpha+1)\int_0^1 g^{1-\alpha}(x)dx$$

with probability one.

Using this last result in (3.19), we see that with probability one, (3.19) is of order $O(n^{-(d_g-1)})$. Thus, (3.17.g) is of the order in n $O(n^{-(d_g-1)})$ w.p.1. By (3.18), we have the order in n of (3.15) is w.p.1.

$$\sum_{g=1}^{e} -(d_g-1) = e-d$$

Since $\Sigma^{(e,d)}$ in (3.14) contains $O(m^d)$ terms of this magnitude, we have w.p.1,

$$(3.24) \quad A_{ed} = n^{-p/2}\, O(m^d)\, O(n^{e-d}) = O(n^{e-p/2}).$$

If $e > p/2$, then from (3.16) there is at least one g such that $p_g=1$ and (3.17.g) vanishes because of (3.9) and (3.4), whence (3.15) also vanishes so that $A_{ed}=0$. We have proved so far that

$$(3.25) \quad A_{ed} = \begin{cases} 0 & \text{if } e > p/2 \\ O(n^{e-p/2}) \text{ w.p.1} & \text{if } e \le p/2 \end{cases}$$

From (3.13) and (3.25) it follows that $\mu_p = o(1)$ w.p.1 for odd p.

In the case when p is even, we need only consider A_{ed} for $e=p/2$ because of (3.25), i.e.,

$$(3.26) \quad \mu_p \sim \sum_{d=2}^{2p} A_{p/2,d} = A \text{ (say)} \quad (\text{w.p.1}).$$

The reasoning which led to (3.25) shows that positive contributions to A are made only when each $p_g=2$, $g=1,\dots,p/2$. This means that each d_g $(g=1,\dots,p/2)$ has to be either 2 or 3. If $d_g=2$, then by (3.23), n times (3.17.g) converges with probability one to

$$(3.27) \quad I(g) = 2\int_0^1 g^{-1}(x)dx$$

If $d_g=3$, we have n^2 times (3.17.g) converging with probability one to

$$(3.28) \quad II(g) = 6\int_0^1 g^{-2}(x)dx - 4\left(\int_0^1 g^{-1}(x)dx\right)^2.$$

If, q of the numbers $d_1,\dots,d_{p/2}$ are 2 and $p/2-q$ are 3, then using (3.27) and (3.28), we have

$$(3.29) \quad \lim_{n\to\infty} n^{p-q} E_n[\psi_2(X_{i_1}, X_{j_1}) \dots \psi_2(X_{i_p}, X_{j_p})|Y] = (I(g))^q (II(g))^{p/2-q}$$

with probability one.

Combining (3.29) with (3.26), (3.14) and the remarks preceding (3.27), we have

$$(3.30) \qquad A = n^{-3p/2} \sum_{q=0}^{p/2} n^q (I(g))^q (II(g))^{p/2-q} \varphi(q) \, \psi(n, q)$$

with probability one, where $\varphi(q)$ is the number of ways of classifying p pairs $(i_1,j_1),\ldots,(i_p,j_p)$ into p/2 sets, q of which have 2 different subscripts, and p/2-q have 3 different subscripts. Clearly,

$$(3.31) \qquad \varphi(q) = \binom{p/2}{q}(p-1)(p-3)\ldots 3.1 = \binom{p/2}{q} \frac{p!}{2^{p/2}\,(p/2)!}$$

By $\psi(n, q)$, we represent the number of ways of choosing $i_1,j_1;i_2,j_2;\ldots;i_p,j_p$, $1 \le i_g < j_g \le m$ so that $i_g=i_{g+1}$; $j_g=j_{g+1}$, $g=1,3,5,9,\ldots,2q-1$; and so that one of the equalities $\left\{\begin{matrix} i_g=i_{g+1} & j_g=i_{g+1} \\ i_g=j_{g+1} & j_g=j_{g+1}\end{matrix}\right\}$ is satisfied for $g=2q+1, 2q+3,\ldots,p-1$.

We can see that $\psi(n, q)$ is given by

$$(3.32) \qquad \psi(n, q) = \binom{m}{2}\binom{m-2}{2}\ldots\binom{m-2q+2}{2}\binom{m-2q}{2} \, 2(m-2q-2)\binom{m-2q-3}{2} \, 2(m-2q-5) \ldots (m+q-3p/2+1)$$

$$= \prod_{j=0}^{q-1}\binom{m-2j}{2} \prod_{j=0}^{(3p/2)-3q-1}(m-2q-j) = (2^{-q}) \prod_{j=0}^{(3p/2)-q-1}(m-j) = (2^{-q})(m^{(3p/2)-q}) \prod_{j=0}^{(3p/2)-q-1}\left(1-\frac{j}{m}\right)$$

Using (3.31) and (3.32) in (3.30), we have

$$(3.33) \qquad A = \frac{p!}{2^{p/2}(p/2)!} \, r^{3p/2} \sum_{q=0}^{p/2}\left(\frac{1}{2r}\right)^q \binom{p/2}{q} (I(g))^q (II(g))^{p/2-q} + o(1)$$

with probability one (recalling that m=rn). Clearly, (3.33) reduces to the simple expression

$$(3.34) \qquad \frac{p!}{2^{p/2}(p/2)!} \, r^p \left\{\int_0^1 g^{-1}(x)dx + 6r\int_0^1 g^{-2}(x)dx - 4r\left(\int_0^1 g^{-1}(x)dx\right)^2\right\}^{p/2}$$

This completes the proof of Theorem 3.1.

4. Asymptotic Distribution of V^2 Statistic.

We can combine the results (3.5), (3.6), and Theorem 3.1 to infer the following

Theorem 4.1 Under the assumptions of Theorems 3.1 and 7.1, the distribution of

$$(4.1)\quad \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n+1}\binom{S_i}{2} - r^2\int_0^1 g^{-1}(x)dx\right)}{r\left[\int_0^1 g^{-1}(x)dx + 6r\int_0^1 g^{-2}(x)dx + 2r^2\int_0^1 g^{-3}(x)dx - r(r+4)\left(\int_0^1 g^{-1}(x)dx\right)^2\right]^{\frac{1}{2}}}$$

approaches the standard normal distribution as n increases.

Since $V^2 = r + 2\left(\frac{1}{n}\sum_{i=1}^{n+1}\binom{S_i}{2}\right)$, we can compute the power of tests based on V^2 using the theorem above. A test of the hypothesis that $G(x) = F(x)$ (the uniform distribution) $(0 \leq x \leq 1)$ based on V^2 would reject this hypothesis whenever V^2 exceeds $C_n(\alpha)$ where α is the desired level of significance. We shall use the following standard notation:

$$(4.2)\qquad \Phi(v) = \left(\frac{1}{\sqrt{2\pi}}\right)\int_v^\infty e^{-(t^2/2)}dt$$

and

$K(\alpha)$ is the number such that $\Phi(K(\alpha)) = \alpha$

Then the above theorem shows that for large n, $C_n(\alpha)$ is approximately equal to

$$(4.3)\qquad \left(\frac{r}{\sqrt{n}}\right)[\sqrt{n}(1+2r) + 2(r+1)K(\alpha)].$$

5. Limiting Conditional Normality of the Run Statistic.

In this section, we consider the limiting conditional distribution of

$$(5.1)\quad H_o(S_1,\ldots,S_{n+1}) = \frac{1}{n+1}\ \text{(the number of } S_1,\ldots,S_{n+1} \text{ which equal zero)}$$

We shall abbreviate $H_o(S_1,\ldots,S_{n+1})$ by H_o. Denote the number of runs of X's and Y's in the combined ordered sample by U_n. It is easily seen that the number of runs of X's is the same as the number of cells containing at least one X, which is $(n+1)(1-H_o)$, and from the definition of U_n, we see that U_n differs from twice the number of runs of X's by at most one. Formally, we have

$$(5.2) \qquad |(U_n/n) - ((n+1)/n)(1-H_o)| \le \frac{1}{n}$$

From (5.2), we see that if H_o is asymptotically normal with mean μ and variance σ^2, U_n/n will be asymptotically normal with mean $2(1-\mu)$ and variance $4\sigma^2$. We shall now examine the distribution of H_o.

Since we have

$$(5.3) \quad P\{S_i=0|Y_1,\ldots,Y_n\} = (1-W_i)^m$$

where W_i is the length of the $i^{\underline{th}}$ spacing, it follows that

$$(5.4) \qquad E_n(H_o|Y) = (1/(n+1)) \sum_{i=1}^{n+1} (1-W_i)^m .$$

It is easy to show that asymptotically $(1/(n+1)) \sum_{i=1}^{n+1} (1-W_i)^m$ has the same distribution as $(1/(n+1)) \sum_{i=1}^{n+1} e^{-nrW_i}$. The asymptotic normality of the latter can be demonstrated by the method employed by Weiss [14], or the generalization of Proschan and Pyke [11] (see Section 7).

From Theorem 2.1 of Section 2, it follows that we need only consider the limiting behavior of the conditional moments of $\sqrt{n}(H_o - E_n(H_o|Y))$. Using the computational scheme which Kitabatake [8] employed to solve the one-sample analogue of this problem, we shall prove the following

<u>Theorem 5.1</u> Let H_o be defined by (5.1) and let $g(x)$ be bounded on [0,1], then

$$(5.5) \quad \lim_{n\to\infty} E_n\{n^{\ell/2}(H_o - E_n(H_o|Y))^\ell|Y\} = \begin{cases} 0 & \text{if } \ell = 1,3,5,\ldots \\ (\ell-1)(\ell-3)\ldots 1C^{\ell/2} & \text{if } \ell = 2,4,6,\ldots \end{cases}$$

with probability one. The constant C is given by

$$(5.6)\quad C = \int_0^1 \frac{g^2(x)}{r+g(x)}\,dx - \int_0^1 \frac{g^2(x)}{2r+g(x)}\,dx - r\Big(\int_0^1 \frac{g^2(x)}{(r+g(x))^2}\,dx\Big)^2$$

Proof:

Let $J_o = (n+1)H_o$. Let $V_i = 1$ if $S_i = 0$, and 0 otherwise. Then,

$$(5.7)\quad J_o = \sum_{i=1}^{n+1} V_i ,$$

and

$$(5.8)\quad E_n\{V_i|Y\} = (1-W_i)^m .$$

Also,

$$(5.9)\quad E_n\{J_o^{(s)}|Y\} = \sum_{nP_s} (1-W_{i_1}-\dots- W_{i_s})^m$$

where

$$J_o^{(s)} = J_o(J_o-1)\ \dots\ (J_o-s+1) \quad \text{if } s > 0$$

$$J_o^{(0)} = 1$$

and $\sum_{nP_s}$ stands for summation over all permutations $(i_1,\dots,i_s)$ of $(n+1)$ integers such that $1 \leq i_j \leq n+1$, $i_j \neq i_k$ if $j \neq k$ $(j,k=1,2,\dots,n+1)$.

We note also that we can write (w.p.1)

$$(5.10)\quad (1-W_{i_1}-\dots- W_{i_s})^m = e^{-r\sum_{j=1}^{s}(nW_{ij})}\Big[1- \frac{r}{2n}\Big(\sum_{j=1}^{s}(nW_{ij})\Big)^2 + 0\Big(\frac{1}{n^2}\Big)\Big].$$

By the binomial expansion, we obtain

$$(5.11)\quad E_n\{n^{\ell+2/2}\Big(\frac{J_o}{n+1} - E_n(H_o|Y)\Big)^{\ell+2}|Y\} =$$

$$= n^{\ell+2/2} E_n\{[\frac{J_o}{n+1} - E_n(H_o|Y)][\frac{J_o}{n+1} - E_n(H_o|Y)]^{\ell+1}|Y\}$$

$$= n^{\ell+2/2}\left[\sum_{j=0}^{\ell+1} (-1)^j \binom{\ell+1}{j} E_n\{(\frac{J_o}{n+1})^{\ell+2-j}|Y\}\{E_n(H_o|Y)\}^j\right.$$

$$\left. - \sum_{j=0}^{\ell+1} (-1)^j \binom{\ell+1}{j} E\{(\frac{J_o}{n+1})^{\ell+1-j}|Y\}\{E_n(H_o|Y)\}^{j+1}\right].$$

We can express J_o^k in terms of factorial powers as

$$(5.12) \qquad J_o^k = \sum_{q=0}^{k} \binom{k}{q} B_q^{(k-q)} J_o^{(k-q)}$$

where $B_r^{(n)}$ is the Stirling number of order n and degree r. From (5.9) and (5.12) we obtain that with probability one

$$(5.13) \quad E_n\{(\frac{J_o}{n+1})^{\ell+2-j}|Y\} = E_n\{\frac{J_o^{(\ell+2-j)}}{(n+1)^{\ell+2-j}}|Y\}$$

$$+ \frac{(\ell+2-j)(\ell+1-j)}{2} \frac{1}{n+1} E_n\{\frac{J_o^{(\ell+1-j)}}{(n+1)^{\ell+1-j}}|Y\} + E_n\{\frac{J_o^{(\ell-j)}}{(n+1)^{\ell-j}}|Y\} O(\frac{1}{n^2}).$$

To evaluate the terms on the right of (5.13), we use (5.9) and (5.10) much computation, and (5.13) again to obtain

$$(5.14) \quad E_n\{\frac{J_o^{(s)}}{(n+1)^s}|Y\} = \sum_{nP_s} \frac{(1-W_{i_1}-\dots-W_{i_s})^m}{(n+1)^s}$$

$$= \sum_{nP_s} \frac{e^{-r(\sum_{j=1}^{s}(nW_{i_j}))}}{(n+1)^s} - \frac{r}{2(n+1)} \sum_{nP_s} (\sum_{j=1}^{s}(nW_{i_j}))^2 \frac{e^{-r(\sum_{j=1}^{s}(nW_{i_j}))}}{(n+1)^s}$$

$$+ \sum_{nP_s} \frac{e^{-r(\sum_{j=1}^{s}(nW_{i_j}))}}{(n+1)^s} O(\frac{1}{n^2}) = \left\{(\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1})(\sum_{nP_{s-1}} \frac{e^{-r\sum_{j=1}^{s-1}(nW_{i_j})}}{(n+1)^{s-1}})\right.$$

$$- \frac{(s-1)}{n+1} \left(\sum_{i=1}^{n+1} \frac{e^{-2rnW_i}}{n+1} \right) \left(\sum_{nP_{s-2}} \frac{e^{-r \sum_{j=i}^{s-2} (nW_{i_j})}}{(n+1)^{s-2}} \right)$$

$$+ \frac{(s-1)(s-2)}{(n+1)^2} \left(\sum_{nP_{s-2}} \frac{e^{-3rnW_{i_1}} e^{-r \sum_{j=2}^{s-2} (nW_{i_j})}}{(n+1)^{s-2}} \right) \Bigg\}$$

$$+ \Bigg\{ - \frac{sr}{2(n+1)} \left(\sum_{i=1}^{n+1} \frac{(nW_i)^2 e^{-rnW_i}}{n+1} \right) \left(\sum_{nP_{s-1}} \frac{e^{-r \sum_{j=1}^{s-1} (nW_{i_j})}}{(n+1)^{s-1}} \right)$$

$$+ \frac{rs(s-1)}{2(n+1)^2} \left(\sum_{nP_{s-1}} \frac{(nW_{i_1})^2 e^{-2rnW_{i_1}} e^{- \sum_{j=2}^{s-1} (nW_{i_j})}}{(n+1)^{s-1}} \right)$$

$$- \frac{rs(s-1)}{2(n+1)} \left(\sum_{i=1}^{n+1} \frac{(nW_i) e^{-rnW_i}}{n+1} \right)^2 \left(\sum_{nP_{s-2}} \frac{e^{-r \sum_{j=1}^{s-2} (nW_{i_j})}}{(n+1)^{s-2}} \right)$$

$$+ \frac{rs(s-1)}{2(n+1)^2} \left(\sum_{i=1}^{n+1} \frac{(nW_i)^2 e^{-2rnW_i}}{n+1} \right) \left(\sum_{nP_{s-2}} \frac{e^{-r \sum_{j=1}^{s-2} (nW_{i_j})}}{(n+1)^{s-2}} \right)$$

$$+ \frac{rs(s-1)(s-2)}{2(n+1)^3} \left(\sum_{nP_{s-2}} \frac{(nW_{i_1})^2 e^{-3rnW_{i_1}} e^{-r \sum_{j=2}^{s-2} (nW_{i_j})}}{(n+1)^{s-2}} \right) \Bigg\}$$

$$+ \left(\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1} \right)^2 \left(\sum_{nP_{s-2}} \frac{e^{-r \sum_{j=1}^{s-2} (nW_{i_j})}}{(n+1)^{s-2}} \right) \quad O\left(\frac{1}{n^2}\right)$$

$$= \left[\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1} - \frac{r}{2(n+1)} \sum_{i=1}^{n+1} \frac{(nW_i)^2 e^{-rnW_i}}{n+1} \right] [E_n\{ \frac{J_o^{s-1}}{(n+1)^{s-1}} |Y\}$$

$$- \frac{(s-1)(s-2)}{2(n+1)} \quad E_n\{ \frac{J_o^{s-2}}{(n+1)^{s-2}} |Y\}] - \frac{(s-1)}{n+1} [r(\sum_{i=1}^{n+1} \frac{(nW_i) e^{-rnW_i}}{n+1})^2$$

$$+ \sum_{i=1}^{n} \frac{e^{-2rnW_i}}{n+1}] E_n\{ \frac{J_o^{s-2}}{(n+1)^{s-2}} |Y\} + O(\frac{1}{n^2}) E_n\{ \frac{J_o^{s-2}}{(n+1)^{s-2}} |Y\}$$

with probability one.

Using (5.14) on the right side of (5.13) with the appropriate values of s and lumping together terms of order $(\frac{1}{n^2})$, we obtain

$$(5.15) \quad E_n\{(\frac{J_o}{n+1})^{\ell+2-j}|Y\} = [\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1}$$

$$- \frac{r}{2(n+1)} (\sum_{i=1}^{n+1} \frac{(nW_i)^2 e^{-rnW_i}}{n+1})] E_n\{(\frac{J_o}{n+1})^{\ell+1-j}|Y\} + \frac{(\ell+1-j)}{n+1} [\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1}$$

$$- \sum_{i=1}^{n+1} \frac{e^{-2rnW_i}}{n+1} - r(\sum_{i=1}^{n+1} \frac{(nW_i) e^{-rnW_i}}{n+1})^2 + O(\frac{1}{n})] E_n\{(\frac{J_o}{n+1})^{\ell-j}|Y\}$$

with probability one.

Putting the result (5.15) into the expansion (5.11) we obtain

$$(5.16) \quad E_n\{n^{\ell+2/2}(\frac{J_o}{n+1} - E_n(H_o|Y))^{\ell+2}|Y\}$$

$$= n^{\ell+2/2} \{ \sum_{j=0}^{\ell+1} (-1)^j \binom{\ell+1}{j} E_n\{(\frac{J_o}{n+1})^{\ell+1-j}|Y\}\{E_n(H_o|Y)\}^j [\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1}$$

$$- \frac{r}{2(n+1)} \sum_{i=1}^{n+1} \frac{(nW_i)^2 e^{-rnW_i}}{n+1} - E_n(H_o|Y)]$$

$$+ \frac{(\ell+1)}{n} \sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} E_n\{(\frac{J_o}{n+1})^{\ell-j}|Y\}\{E_n(H_o|Y)\}^j [\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1} - \sum_{i=1}^{n+1} \frac{e^{-2rnW_i}}{n+1}$$

$$-r(\sum_{i=1}^{n+1} \frac{(nW_i) e^{-rnW_i}}{n+1})^2 + o(\tfrac{1}{n})]\} = \sqrt{n} [\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1}$$

$$- \frac{r}{2(n+1)} \sum_{i=1}^{n} \frac{(nW_i)^2 e^{-rnW_i}}{n+1} - E_n(H_o|Y)] E_n\{n^{\ell+1/2}(\frac{J_o}{n+1} - E_n(H_o|Y))^{\ell+1}|Y\}$$

$$+ (\ell+1) E_n\{n^{\ell/2}(\frac{J_o}{n+1} - E_n(H_o|Y))^{\ell}|Y\}[\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1} - \sum_{i=1}^{n+1} \frac{e^{-2rnW_i}}{n+1}$$

$$- r(\sum_{i=1}^{n+1} \frac{(nW_i) e^{-rnW_i}}{n+1})^2 + o(\tfrac{1}{n})] = o(\frac{1}{n^{3/2}}) E_n\{n^{\ell+1/2}(\frac{J_o}{n+1} - E_n(H_o|Y))^{\ell+1}|Y\}$$

$$+ (\ell+1) E_n\{n^{\ell/2}(\frac{J_o}{n+1} - E_n(H_o|Y))^{\ell}|Y\}[\sum_{i=1}^{n+1} \frac{e^{-rnW_i}}{n+1} - \sum_{i=1}^{n+1} \frac{e^{-2rnW_i}}{n+1}$$

$$- r(\sum_{i=1}^{n+1} \frac{(nW_i) e^{-rnW_i}}{n+1})^2 + o(\tfrac{1}{n})]$$

with probability one.

By an inductive argument, we obtain from (5.16) that w.p.1

$$(5.17)\quad E_n\{n^{\ell}\,(H_o-E_n(H_o|Y))^{2\ell}|Y\} = (2\ell-1)(2\ell-3)\ldots5.3.1\,[\sum_{i=1}^{n+1}\frac{e^{-rnW_i}}{n+1}$$

$$-\sum_{i=1}^{n+1}\frac{e^{-2rnW_i}}{n+1} - r(\sum_{i=1}^{n+1}\frac{(nW_i)\,e^{-rnW_i}}{n+1})^2]^{\ell} + 0(\tfrac{1}{n})$$

and

$$E_n\{n^{\ell+\frac{1}{2}}(H_o-E_n(H_o|Y))^{2\ell+1}|Y\} = 0(\frac{1}{n^{3/2}})$$

Using the Weiss convergence result, Theorem 3.2, we have

$$(5.18)\quad \lim_{n\to\infty}\,[\sum_{i=1}^{n+1}\frac{e^{-rnW_i}}{n+1} - \sum_{i=1}^{n+1}\frac{e^{-2rnW_i}}{n+1} - r(\sum_{i=1}^{n+1}\frac{(nW_i)\,e^{-rnW_i}}{n+1})^2]$$

$$= \int_0^1 \frac{g^2(x)}{r+g(x)}\,dx - \int_0^1 \frac{g^2(x)}{2r+g(x)}\,dx - r(\int_0^1 \frac{g^2(x)}{(r+g(x))^2}\,dx)^2$$

with probability one.

Taking limits in (5.17) and using (5.18) yields the desired result (5.5). Thus Theorem 5.1 is proved.

6. Asymptotic Distribution of the Run Statistic.

Now we can put together the remark following (5.2) with the results of Theorem 2.1, Theorem 5.1 and Theorem 7.1, to obtain explicitly the limiting distribution of U_n, the number of runs in the combined sample of X's and Y's.

Theorem 6.1 Under the assumptions of Theorems 5.1 and 7.1, the distribution of

$$(6.1)\quad \frac{\sqrt{n}\,(\frac{1}{n}\,U_n - 2\int_0^1 \frac{rg(x)}{r+g(x)}dx)}{2[\int_0^1 \frac{rg(x)}{r+g(x)}dx - 2r^2\int_0^1 \frac{g^2(x)}{(r+g(x))^3}dx - r(\int_0^1 \frac{g^2(x)}{(r+g(x))^2}dx)^2 - r^4(\int_0^1 \frac{g\;(x)}{(r+g(x))^2}dx)^2]^{\frac{1}{2}}}$$

approaches the standard normal as n increases.

This is the same as the expression derived by Wolfowitz [16] who used a somewhat controversial method of proof.

Using the theorem, we can set up a test of the hypothesis that $G(x) = F(x)$ (the uniform distribution) based on U_n and having size of approximately α for large n. Letting $\Phi(v)$ and $K(\alpha)$ be defined by (4.2), the test based on U_n will reject the hypothesis of equality whenever U_n/n is less than

$$(6.2)\qquad (2r/(1+r))[1 - (K(\alpha)\sqrt{1+r}/(1+r)\sqrt{n})].$$

7. Distributions of Functions of Sample Spacings.

In the proofs of normality in Sections 3 and 5 and in the computations of power in Sections 4 and 6, we used the asymptotic normality of certain functions of sample spacings. In particular, we used the

Theorem 7.1 If $g(x)$ is bounded from above and below away from 0 on [0,1], then as n increases, the marginal distributions of

$$(7.1)\qquad \frac{\sqrt{n}\,\left(\frac{1}{n}\sum_{i=1}^{n+1}(nW_i)^2 - 2\int_0^1 \frac{1}{g(x)}\,dx\right)}{2\left[2\int_0^1 \frac{1}{g^3(x)}\,dx - \left(\int_0^1 \frac{1}{g(x)}\,dx\right)^2\right]^{\frac{1}{2}}}$$

and

$$(7.2)\qquad \frac{\sqrt{n}\,\left(\frac{1}{n}\sum_{i=1}^{n+1}(1-W_i)^m - \int_0^1 \frac{g^2(x)}{r+g(x)}\,dx\right)}{\left[\int_0^1 \frac{g^2(x)}{2r+g(x)}dx + \int_0^1 \frac{rg(x)}{r+g(x)}dx - \int_0^1 \frac{g^2(x)}{r+g(x)}dx - 2r^2\int_0^1 \frac{g^2(x)}{(r+g(x))^3} - r^4\left(\int_0^1 \frac{g\ (x)}{(r+g(x))^2}\right)^2\right]^{\frac{1}{2}}}$$

each approach the standard normal distribution.

Proof:

In [14], Weiss announced the result (3.6) which specializes to (7.1). The derivation given in [14] holds strictly only for $g(x)$ a step function. When $g(x)$ is continuous and has a continuous derivative, the results announced by Proschan and Pyke [11], imply (7.1). The result (7.1) can be derived using the method employed below to derive (7.2).

The result (7.2) also would follow from the results announced in [11]. The method used there is based on a series expansion of the functions involved (Pyke-personal communication). We shall illustrate the method by deriving (7.2) when $g(x)$ has a bounded second derivative, and $g(x)$ is bounded away from zero.

To obtain (7.2), we observe first that $(\sqrt{n}/n)\Sigma(1-W_i)^{nr}$ has the same distribution as $(\sqrt{n}/n)\Sigma e^{-nrW_i}$. Denote $G(Z_i) - G(Z_{i-1})$ by U_i $(i=1,\ldots,n+1)$. Note that

$$(7.3) \qquad G(Z_i) = \sum_{j=1}^{i} U_j \qquad i=1,\ldots,n+1.$$

$U_1,\ldots,U_{n+1}$ are distributed as sample spacings based on n observations from a uniform distribution. Also, if $V_1,\ldots,V_{n+1}$ are exponentially distributed $(f(v) = e^{-v})$, and if $T_n = V_1 +\ldots+ V_{n+1}$, then $U_1,\ldots,U_{n+1}$ and $\frac{V_1}{T_n},\ldots,\frac{V_{n+1}}{T_n}$ have the same joint distribution.

By the mean value theorem,

$$(7.4) \quad U_i = g(\hat{Z}_i)W_i = gG^{-1}(\tfrac{i}{n})W_i + (g(\hat{Z}_i) - gG^{-1}(\tfrac{i}{n}))W_i \quad (Z_{i-1} < \hat{Z}_i < Z_i).$$

Thus,

$$W_i = U_i/gG^{-1}(\tfrac{i}{n}) + W_i(gG^{-1}(\tfrac{i}{n}) - g(\hat{Z}_i))/gG^{-1}(\tfrac{i}{n})$$

$$(7.5) \quad = U_i/gG^{-1}(\tfrac{i}{n}) + U_i(gG^{-1}(\tfrac{i}{n}) - g(\hat{Z}_i))/g^2G^{-1}(\tfrac{i}{n})$$

$$+ W_i^2(gG^{-1}(\tfrac{i}{n}) - g(\hat{Z}_i))^2/g^2G^{-1}(\tfrac{i}{n}).$$

Also,

$$(7.6)\quad g(\hat{Z}_i) - gG^{-1}(\tfrac{i}{n}) = (G(Z_i) - \tfrac{i}{n}) \frac{d}{dx} gG^{-1}(x)\Big|_{x=\frac{i}{n}}$$

$$- (G(Z_i) - G(\hat{Z}_i)) \frac{d}{dx} gG^{-1}(x)\Big|_{x=\frac{i}{n}}$$

$$+ (G(\hat{Z}_i) - \tfrac{i}{n})^2 \frac{d^2}{dx^2} gG^{-1}(x)\Big|_{x=\theta} \qquad \frac{i}{n} < \theta < G(\hat{Z}_i).$$

Note that

$$(7.7)\quad \frac{d}{dx} g(G^{-1}(x)) = \left(\frac{d}{du} g(u)\Big|_{u=G^{-1}(x)}\right) \frac{1}{g(G^{-1}(x))} .$$

We shall denote $gG^{-1}(\frac{i}{n})$ by g_i and $\frac{d}{du} g(u)\Big|_{u=G^{-1}(x)}$ by g_i'.

Now we shall expand $(\sqrt{n}/n)\Sigma e^{-rnW_i}$ around $(\sqrt{n}/n)\Sigma e^{-rnU_i/g_i}$ in a Taylor series with remainder term based on the second derivative, using (7.5), (7.6) and (7.7):

$$(7.8)\quad \frac{\sqrt{n}}{n} \Sigma e^{-rnW_i} = \frac{\sqrt{n}}{n} \Sigma (e^{-rnU_i/g_i} + r\frac{g_i'}{g_i^3} (G(Z_i) - \tfrac{i}{n})(nU_i) e^{-rnU_i/g_i})$$

$$+ \frac{r\sqrt{n}}{n} \Sigma\, (nU_i) e^{-rnU_i/g_i} [\, \frac{g_i'}{g_i^3} (G(\hat{Z}_i) - G(Z_i)) + \frac{\frac{d^2}{dx^2} gG^{-1}(x)\Big|_{x=\theta}}{g_i^2} (G(\hat{Z}_i) - \tfrac{i}{n})^2]$$

$$- r \frac{\sqrt{n}}{n} \Sigma \frac{(g_i - g(\hat{Z}_i))^2}{g_i^2} (nW_i)^2 [\, \frac{e^{-rnU_i/g_i}}{n} - r\, e^{-rnU_i/gG^{-1}(\theta)} \,] \quad (\tfrac{i}{n} < \theta < G(\hat{Z}_i)).$$

We now observe that of the three summations on the right side of (7.8), the latter two converge stochastically to zero. This observation follows from the assumptions about $g(x)$ and its derivatives, the fact that $\frac{1}{n} \Sigma\, (nU_i)^k e^{-rnU_i}$ is uniformly bounded for all n, the fact that $G(Z_i) - G(\hat{Z}_i) < U_i$, and the

Glivenko-Cantelli Lemma which says that $n^{\frac{1}{2}-\delta} \sup_i (G(Z_i)- \frac{i}{n})$ converges stochastically to zero. (In the terms in question, $(G(Z_i)- \frac{i}{n})$ appears as a square.) Using the Slutsky Proposition (Cramèr [2], p. 255), we can say that $(\sqrt{n}/n)\Sigma e^{-rnW_i}$ has the same limiting distribution as

$$(7.9) \quad \frac{\sqrt{n}}{n} \Sigma\, e^{-rnU_i/g_i} [1 + r(g_i'/g_i^3)\, (G(Z_i)- \frac{i}{n})\, (nU_i)].$$

Using the remark following (7.3), we see that (7.9) has the same limiting distribution as

$$(7.10) \quad \frac{\sqrt{n}}{n} \Sigma\, e^{-rnV_i/T_n g_i} [1+r(\frac{g_i'}{g_i^3})\, (G(Z_i)- \frac{i}{n})\, (nV_i/T_n)]$$

where the V_i are independent exponential random variables ($i=1,\ldots,n+1$), and $T_n=\Sigma V_i$. We now examine separately each of the terms in (7.10). Expanding the first term around $(\sqrt{n}/n)\Sigma e^{-rnV_i/g_i}$ we obtain

$$(7.11) \quad \frac{\sqrt{n}}{n} \Sigma\, e^{-rnV_i/T_n g_i} = \frac{\sqrt{n}}{n} \Sigma\, e^{-rV_i/g_i} [1+(r/g_i)\, V_i(1- \frac{n}{T_n})]$$

$$+ \frac{\sqrt{n}}{n} \Sigma\, (rV_i/g_i)^2\, (1- \frac{n}{T_n})^2\, e^{-r\overline{V}_i/g_i}$$

where $V_i/T_n < \overline{V}_i < V_i$. Noting that $n^{\frac{1}{2}-\delta}(1- \frac{n}{T_n})$ approaches zero stochastically, that T_n/n approaches unity stochastically, that

$$(7.12) \quad 1- \frac{n}{T_n} = [\frac{1}{n} \Sigma\, (V_i-1)] \big/ T_n/n$$

and the boundedness of xe^{-x} again, and the fact that

$$(7.13) \quad \frac{1}{n} \Sigma\, \frac{r}{g_i}\, V_i\, e^{-rV_i/g_i} \xrightarrow{P} \int_0^1 \frac{rg^2(x)}{(r+g(x))^2}\, dx$$

as n increases, we can again use the Slutsky Proposition to conclude that $(\sqrt{n}/n)\Sigma\, e^{-rnV_i/T_n g_i}$ has the same distribution as

$$(7.14) \quad (\sqrt{n}/n)\ \Sigma\ (e^{-rV_i/g_i} + (V_i-1)\int_0^1 \frac{rg^2(x)}{(r+g(x))^2} \quad).$$

The second term in (7.10) is a little more complicated, but using (7.3) to express $(G(Z_i)-\frac{i}{n})$ as a function of $V_1,\ldots,V_{n+1}$, and (7.12), and the Glivenko-Cantelli Lemma and the fact that $n^{\frac{1}{2}-\delta}\ (1-\frac{n}{T_n})$ approaches zero stochastically, along with the Slutsky Theorem, we conclude that the limiting distribution of $(\sqrt{n}/n)\ \Sigma\ r(g_i'/g_i^3)(G(Z_i)-\frac{i}{n})(nV_i/T_n)\ e^{-rnV_i/T_n g_i}$ is the same as that of

$$(7.15) \quad r(\sqrt{n}/n) \sum_{i=1}^{n+1} (g_i'/g_i^2)\ V_i\ e^{-rV_i/g_i}\ (1/n) \sum_{j=1}^{i} (V_j-1)$$

$$- (V_i-1)\ (1/n) \sum_{j=1}^{n+1} (j/n)(g_j'/g_j^2)\ V_j\ e^{-rV_j/g_j}$$

Since

$$\frac{1}{n} \sum_{j=1}^{n+1} \frac{j}{n}\ (g_j'/g_j^2)\ V_j\ e^{-rV_j/g_j} \xrightarrow{P} \int_0^1 \frac{G(x)g'(x)}{(r+g(x))^2}\, dx$$

(7.16) and

$$\frac{\sqrt{n}}{n} \sum_{j=1}^{n+1} (V_j-1) \left\{ \frac{1}{n} \sum_{i=j}^{n+1} (g_i'/g_i^2)\ V_i\ e^{-rV_i/g_i} - \int_{G^{-1}(j/n)}^{1} \frac{g'(x)}{(r+g(x))^2}\, dx \right\} \xrightarrow{P} 0$$

as n increases, by re-arranging the order of the first two summations in (7.15) and using Slutsky's Theorem, we find that (7.15) has the same limiting distribution as

$$(7.17) \quad r(\sqrt{n}/n)\ \Sigma\ (V_i-1)\ \left[\int_{G^{-1}(j/n)}^{1} \frac{g'(x)}{(r+g(x))^2} dx - \int_0^1 \frac{G(x)g'(x)}{(r+g(x))^2} dx \right]$$

Integrating the expression in brackets in (7.17), we observe that (7.17) simplifies to

$$(7.18) \quad r(\sqrt{n}/n)\ \Sigma\ (V_i-1)\ \left(\frac{1}{r+g_i} - \int_0^1 \frac{g(x)}{r+g(x)}\, dx \right).$$

Putting together (7.18) and (7.14), we conclude that $(\sqrt{n}/n)\ \Sigma\ e^{-rnW_i}$ has the same limiting distribution as

$$(7.19) \quad (\sqrt{n}/n)\ \Sigma\ e^{-rV_i/g_i} + r(V_i-1)\ \left(\frac{1}{r+g_i} - \int_0^1 \frac{rg(x)}{(r+g(x))^2} \right).$$

It now involves only a simple computation to verify the mean and variance given in (7.2), and limiting normality follows from the standard central limit theorem applied to (7.19) which is a sum of independent random variables.

8. A Comparison of Limiting Power.

As an application of the results of Sections 4 and 6, we shall compute the limiting power of the V^2 test and the run test against sequences of alternatives approaching the uniform distribution, and we shall obtain an expression for the relative efficiency of the two tests. We consider a sequence of densities $g_n(x)$ given by

$$(8.1) \quad g_n(x) = 1 + (c/n^{\frac{1}{4}})\ h(x)$$

where $c > 0$ and we have

$$(8.2) \quad \int_0^1 h(x)dx = 0\ ; \qquad |h(x)| < B < \infty\ ; \qquad ch(x) > -1.$$

We define K_α and $\Phi(v)$ as in (4.2).

We shall use the results of Noether (see Fraser [6], pp. 272-273) to compute limiting power and efficiency with respect to the sequence $g_n(x)$.

We start by considering $g(x)$ given by

$$(8.3) \qquad g(x) = 1 + ch(x)$$

where (8.2) is satisfied.

Since the V^2 statistic has the same behavior as $V_n^* = (1/n) \sum_{i=1}^{n+1} \binom{S_i}{2}$, we can use Theorem 4.1 to evaluate its limiting power. We find easily that the mean of V_n^* is

$$(8.4) \qquad r^2 \int_0^1 g^{-1}(x)dx = r^2\left(1+c^2 \int_0^1 h^2(x)(1+ch'(x))^{-1}dx \right).$$

Thus, m=2 in Noether's Theorem. Further, we see that if we denote the denominator of (4.1) by $\sigma_c(V_n^*)$, then

$$(8.5) \qquad \lim_{n \to \infty} \left[\left(\frac{d^2}{dc^2} E_c \left.(V_n^*)\right|_{c=0} \right)\Big/ n^{2\nu}(\sigma_o(V_n^*)/n^{\frac{1}{2}})\right] = [2r/(r+1)] \int_0^1 h^2(x)dx$$

when $\nu=\frac{1}{4}$.

It is easily verified that the remaining conditions of Noether's Theorem are satisfied and that the limiting power of the V^2 test against the sequence (8.1) is

$$(8.6) \qquad \Phi(K_\alpha - (c^2r/(r+1)) \int_0^1 h^2(x)dx).$$

To find the limiting power of the U_n test, we observe that an equivalent test is the one which rejects the null hypothesis whenever $U_n^* = [(2r/(r+1)) - U_n/n]$ is large. From (6.1) we see that the expected value of U_n^* under (8.3) is

$$(8.7)\quad (2r^2/(r+1))\int_0^1 [(1-g(x))/(r+g(x))]dx$$

$$= [2r^2c^2/(1+r)^2]\int_0^1 h^2(x)(1+r+ch(x))^{-1}dx$$

Again it is easily verified that m=2. Denoting the denominator of (6.1) by $\sigma_c(U_n^*)$, we have

$$(8.8)\quad \lim_{n\to\infty} [(\frac{d^2}{dc^2} E_c\ (U_n^*)\Big|_{c=0})/n^{2\nu}(\sigma_o(U_n^*)n^{\frac{1}{2}})] = (2r/(r+1)^{3/2})\int_0^1 h^2(x)dx$$

when $\nu=\frac{1}{4}$.

We can easily show that the remaining conditions of the theorem are true wo that the limiting power of the U_n test is given by

$$(8.9)\quad \Phi(K_\alpha - (c^2r/(r+1)^{3/2})\int_0^1 h^2(x)dx\ ).$$

It is then easily verified that the efficiency of the run test relative to the V^2 test is 1/(r+1). Thus as r, the ratio m/n, increases the relative efficiency of the run test decreases to zero.

REFERENCES

[1] Blum, J. R. and Weiss, L. (1957). Consistency of certain two-sample tests, Ann. Math. Statist. 28 277-280.

[2] Cramer, H. (1950). Mathematical Methods of Statistics. Princeton Univ. Press.

[3] Daniels, H. E. (1944). The relation between measures of correlation in the universe of sample permutations. Biometrika. 33 129-135.

[4] David, F. N. (1950). Two combinatorial tests of whether a sample has come from a given population. Biometrika. 37 97-110.

[5] Dixon, W. J. (1940). A criterion for testing the hypothesis that two samples are from the same population. Ann. Math. Statist. 11 199-204.

[6] Fraser, D. A. S. (1957). Nonparametric Methods in Statistics. Wiley, New York.

[7] Hoeffding, W. (1951). A combinatorial central limit theorem. Ann. Math. Statist. 22 558-566.

[8] Kitabatake, S. (1958). A remark on a non-parametric test. Math. Japon. 5 45-49.

[9] Okamoto, M. (1952). On a nonparametric test. Osaka Math. Jour. 4 71-85.

[10] Okamoto, M. (1952). Some combinatorial tests of goodness of fit. Osaka Math. Jour. 4 215-228.

[11] Proschan, F. and Pyke, R. (1962). General limit theorems for spacings. Ann. Math. Statist. 33 307.

[12] Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population. Ann. Math. Statist. 11 147-162.

REFERENCES

[13] Weiss, L. (1955). The stochastic convergence of a function of sample successive differences. Ann. Math. Statist. 26 532-536.

[14] Weiss, L. (1957). The asymptotic power of certain tests of fit based on sample spacings. Ann. Math. Statist. 28 783-786.

[15] Wilks, S. S. (1960). A combinatorial test for the problem of two samples from continuous distributions. Proc. Fourth Berkeley Symp. Math. Stat. Prob. Univ. of California Press.

[16] Wolfowitz, J. (1949). Nonparametric statistical inference. Proc. Berkeley Symp. Math Stat. Prob. Univ. of California Press.